UNIVERSITY OF TWENTE.

SURFmap

A Network Monitoring Tool Based on the Google Maps API

User Manual

Author Rick Hofstede

Address University of Twente, The Netherlands

Date December 16, 2013

Version 3.2.1

Contents

1 Introduction **3**

2 Installation **4**

2.1 Installation Requirements 4

3 Configuration **5**

4 Using SURFmap **7**

5 Troubleshooting **12**

1 Introduction

SURFmap is a network monitoring tool based on the Google Maps API. It adds a geographical dimension to network traffic, captured by using Cisco's NetFlow [1] or the newer IPFIX [2]. SURFmap runs as a plugin inside NfSen [3] and reads the data captured and stored by NfSen. Due to the fact that SURFmap needs to know geographical locations of hosts, IP addresses need to be converted to geographical locations. SURFmap supports two so-called "geolocation" databases, namely IP2Location [4] and MaxMind [5]. IP2Location offers a commercial and offline database solution, which results in a fast and unlimited geolocation procedure. MaxMind offers besides a commercial offline database solution also a free offline database solution. Although the accuracy of this free service is not as high as the accuracy of the commercial service, it offers roughly the same performance (since it is stored on your own machine). We do therefore strongly recommend you to use either a MaxMind solution, or one of the commercial IP2Location products. All their products containing "Country", "Region" and "City" fields are supported by SURFmap.

SURFmap has been optimized and tested for use in Mozilla FireFox (3+), Apple Safari (4+), Google Chrome (12+) and Microsoft Internet Explorer (7+).

The SURFmap source code (and this manual) are available through the SURFmap project's Web page on Sourceforge. This page is reachable by the following URLs:

- SURFmap project main page: `http://surfmap.sf.net`
- SURFmap project download page: `http://sourceforge.net/projects/surfmap/files/`

The work on SURFmap has been supported by the following publications:

1. Rick Hofstede, Tiago Fioreze. *SURFmap: A Network Monitoring Tool Based on the Google Maps API*, Application session proceedings of the 11th IFIP/IEEE International Symposium on Integrated Network Management (IM 2009), 1-5 June 2009, Long Island, New York, USA, ISBN 978-1-4244-3487-9, pp. 676-690.
2. Rick Hofstede, Anna Sperotto, Tiago Fioreze, Aiko Pras. *The Network Data Handling War: MySQL vs. NfDump*, Proceedings of the 16th EUNICE Open European Summer School 2010 (EUNICE 2010), 28-30 June 2010, Trondheim, Norway. Lecture Notes in Computer Science, Vol. 6164, ISSN 0302-9743 ISBN 978-3-642-13970-3, pp. 167-176.

The following two chapters cover some details on SURFmap's installation and configuration process. Chapter 4 describes the main concepts of SURFmap, while Chapter 5 closes this manual by providing some troubleshooting information.

2 Installation

This chapter outlines details on the installation of SURFmap, which have not been included in the readme file (*readme.txt*). We refer to this file for the regular installation instructions.

2.1 Installation Requirements

In order to achieve the best experience when using SURFmap, the following components should be installed:

- NfSen
- PHP 5.2.4 or newer, together with the following modules: *mbstring*, *cURL*, *PDO SQLite3*

These requirements translate to the following packages:

	Debian/Ubuntu	**RHEL/CentOS**	**FreeBSD**
PHP *cURL* module	`php5-curl`	`php-curl`	
PHP *mbstring* module		`php-mbstring`	
PHP *PDO SQLite3* module	`php5-sqlite`	`php-pdo`	

3 Configuration

The SURFmap configuration (by means of *config.php*) comes with a set of preconfigured settings, which may be adjusted according to your setup. This is especially the case when you installed SURFmap without the use of the installation script, since the script will determine the appropriate values for all essential settings. This chapter will discuss the settings that require some more explanations, besides the short descriptions included in the configuration file.

`default_flow_record_count` – Represents the number of flow records to be used in the flow filter. When *List Flows* is selected, it means that the first N flow records are selected for visualization. In case *Stat TopN* is selected, SURFmap will select the top N aggregated records from the flow data set.

`default_query_type` – Indicates whether SURFmap should use *List Flows* or *Stat TopN*.

`default_query_type_stat_order` – In case *Stat TopN* has been selected as the default query type (see `$config['default_query_type']`), this setting specifies whether the top statistics should be based on flows, packets or bytes.

`resolve_hostnames` – DNS hostnames belonging to IP addresses shown in marker and line information windows, can be resolved. This setting enables or disables this functionality. Although SURFmap is designed to perform DNS hostname lookups in a conservative manner, you may consider disabling DNS hostname resolving in case you don't or cannot make too many request to your DNS server.

`order_flow_records_by_start_time` – If flow data from multiple flow exporters is accumulated in a single NfSen profile, you may consider to use start time sorting for retrieving the so-called "heavy-hitters".

`nfsen_config` – This setting is essential for getting SURFmap to run (it is set by the installation script in case it was used for installing SURFmap). It contains the path to NfSen's configuration file, in which most required file paths are specified.

`nfsen_default_sources` – When multiple data sources are defined within NfSen, SURFmap will always retrieve flow data from all sources if no default source has been specified. This setting allows to specify which subset of sources should be selected for visualization. In case the specified sources cannot be found within the selected profile (which is specified per NfSen plugin in *nfsen.conf*), all available sources are selected.

`internal_domains` – Internal domains can be used in two ways. First and most important, it allows to specify location names for IP ranges that cannot be geolocated, such as (private) IP address ranges behind a NAT. Second, geolocation data can be overridden, because of inaccuracies in the geolocation database, for example.

IP address ranges that are considered 'internal domains' should be specified as prefixes in *nfdump* filter notation [6]. Country, region and city names must be specified for each range. Multiple entries can be specified, as long as the correct syntax is used (it is an array, so elements should be comma-separated).

`demo_mode` - If enabled, SURFmap will enter a special mode for demoing purposes (in contrast to data analysis purposes in an interactive manner). As such, it will not show the menu panel and legend, and it will select and click a visible line on the Google Maps map randomly.

All available settings in *config.php* come with a short description of their meaning, together with default or example values. In case an option is used erroneously (*e.g.* the used syntax is

incorrect), an appropriate error message should be shown when loading the SURFmap frontend in your Web browser. If you still encounter problems while setting up SURFmap, while check out Chapter 5 of this manual.

4 Using SURFmap

This chapter will guide you through a number of key concepts of SURFmap. This will help you understand why and how network data is visualized by SURFmap. Please keep in mind to run SURFmap always from NfSen (*i.e.* start the plugin from NfSen's plugin page), and not as a standalone application. Before covering the key concepts, the following terms and features need to be explained first:

- **NfSen options** - NfSen distinguishes the following two data gathering options: 1) *List flows*, and 2) *Stat TopN*. The first option is just an ordinary listing of the first N flows in the selected time period, eventually ordered and selected based on a filter. On the other hand, *Stat TopN* provides statistics based on the flows in the selected data set. For instance, the top 20 of flows ordered by bytes could be queried. *Stat TopN* statistics can be based on flows, packets or bytes. Besides that, these statistics can be limited to the top *N*.
- **Zoom levels** - Besides of the zoom levels offered by the Google Maps API, SURFmap offers another four zoom levels:

  1. *Country* zoom level
  2. *Region* zoom level
  3. *City* zoom level
  4. *Host* zoom level.

  Each of these zoom levels provides network data at another level of abstraction, where the country zoom level is the least detailed one, and the host zoom level the most detailed one.
- **Line colors** - The colors of the lines are based on their weight. This weight can be based on various network parameters, namely the amount of flows, packets or octets. When NfSen's *List Flows* option is selected, a flow's weight is based on the amount of flows represented by a line. Otherwise, it is based on the *Stat TopN* selection field (*i.e.* either flows, packets or octets). Depending on whether flows, packets or octets are selected as the basis of *Stat TopN*, the line colors are calculated based on a four-color classification scheme.
- **Green marker** - A green marker indicates that there is traffic of which both the source and destination reside *within* the marker. For example, when there is traffic between Amsterdam and Enschede in The Netherlands, this traffic will be aggregated into a single country marker at the *Country* level. This marker will be green.
- **GeoFilter** - SURFmap supports the *GeoFilter* feature starting from v2.3, next to the default *nfdump* filtering. It provides a simple post-processing step for filtering flow data based on geographical metrics (in contrast to the network-based metrics provided by *nfdump*). This means that the *GeoFilter* is always applied *after* the *nfdump* filter. The *GeoFilter* language uses a grammar similar to *nfdump*'s query language and consists of the following operator types:

  – **Logical operators**: **not**, **and**, **or**
  – **Origin operators**: **src**, **dst** (**any** if not explicitly specified)
  – **Location operators**: **country**, **region**, **city**, **ctry**, **rgn**, **cty**.

  Note that all keywords and operands are case-insensitive. Some example queries:

```
src ctry NL
(not src ctry CZ)
src ctry NL and (dst rgn GELDERLAND) and (cty Enschede or CTY Hengelo)
```

Depending on the amount of flow data available on your NetFlow collector, SURFmap is subject to a certain query time. As soon as a query command is fired from the Web interface, *nfdump* [6] will be called in order to do the actual querying. These queries cannot be stopped from the Web interface anymore. The only way to do that is by killing the query process on your system. Although the Web interface will be (temporarily) disabled when you submit a query, it is possible to submit more than one query at once, for example by reloading the Web page. Depending on your system performance and the query impact, you can harm your system severely when doing this. Although a warning message will be shown for potentially *heavy* queries, you should be aware of the performance implications when submitting a query.

The remainder of this chapter consists of a description of various screenshots of SURFmap.

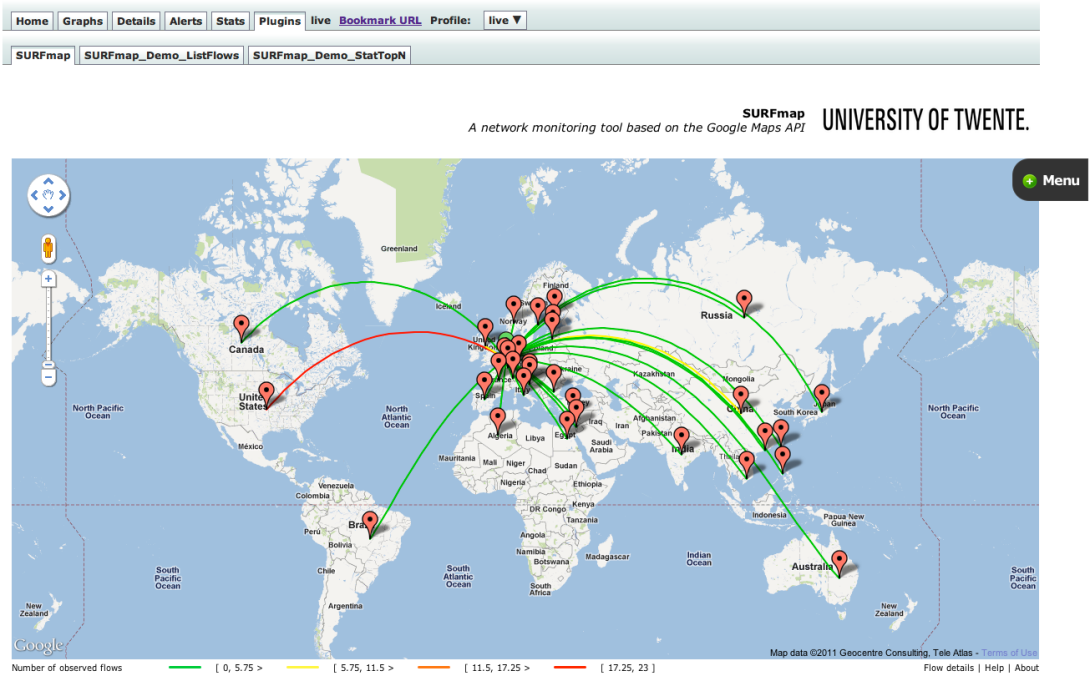

Figure 1: SURFmap at *Country* zoom level

Figure 1 shows a screenshot of the main screen of SURFmap. The selected NfSen option is *List Flows*, which means that the used data set is a pure listing of flows. As a result of this and due to the fact that the screenshot is taken at the *Country* zoom level, the colors of the lines are based on the amount of flows to each country (in the selected time window).

Earlier in this chapter it was explained the a line's color on the map depends on the weight of a line. In Figure 1, a green line is shown between the Netherlands and Canada, while a red line is shown between the Netherlands and the United States. As a consequence, we can conclude that more flows are exchanged between the Netherlands and the United States, than between the Netherlands and Canada.

The actual line color classification of the current SURFmap session and zoom level can be found in the legend below the Google Maps map. In the indicated case, it can be concluded that at least 18 and at most 23 flows were exchanged between the Netherlands and the United States.

Finally, the marker in the Netherlands is colored green. This is a result of the fact that there will be at least one flow of which both the source and destination reside within the Netherlands.

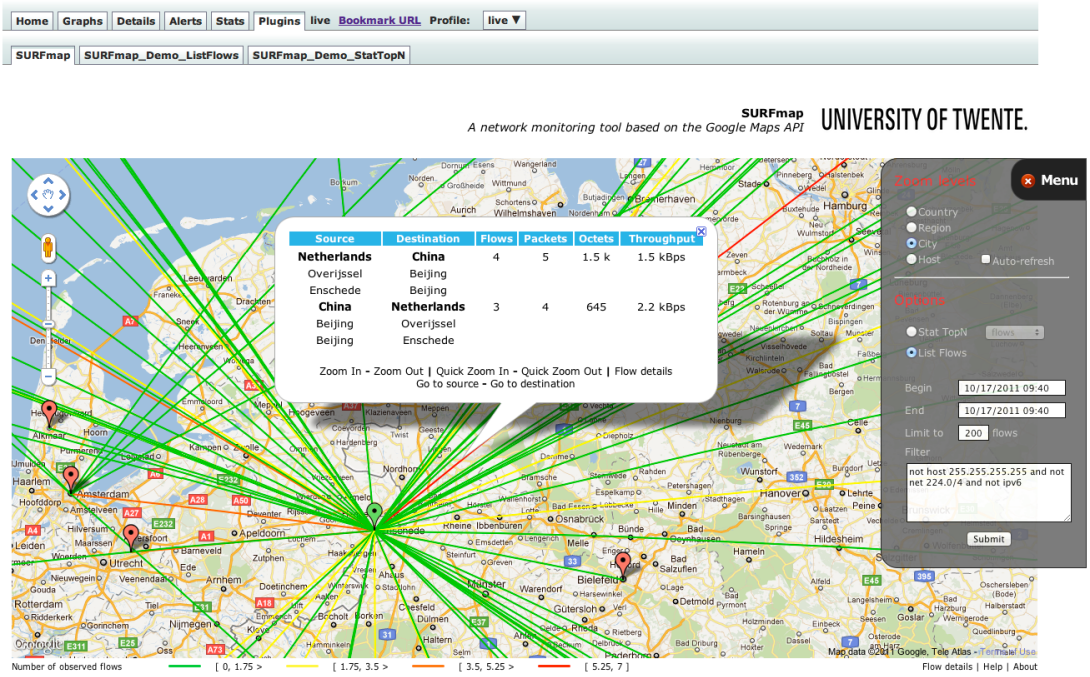

Figure 2: Line information window at *Region* zoom level

Figure 2 shows SURFmap again NfSen's *List Flows* option, but now at the *City* zoom level. The map has been zoomed in to the city of Enschede (the Netherlands), since that city has been configured as the map center in SURFmap's configuration file in the current setup. The map shows one read line, which should represent (according to the legend below the map) 6-7 flows. After clicking on that line, the information window shows that there were indeed 7 flows between Enschede (the Netherlands) and Beijing (China):

- 4 flows from Enschede to Beijing, in which 1.5 kB, spread over 5 packets, were transmitted
- 3 flows from Beijing to Enschede, in which 645 B, spread over 4 packets, were transmitted.

Various buttons/links can be found at the bottom of the information window:

- **Zoom In/Out** Zooms in/out for one Google Maps zoom level. Please note that Google Maps zoom levels are taking *smaller* steps than SURFmap zoom levels
- **Quick Zoom In/Out** Zooms in/out for one SURFmap zoom level. Please note that SURFmap zoom levels are taking *larger* steps than Google Maps zoom levels
- **Zoom In/Out** Shows/hides a table showing the NetFlow data of all visualized flows
- **Zoom In/Out** Jumps to the source/destination of the currently selected line.

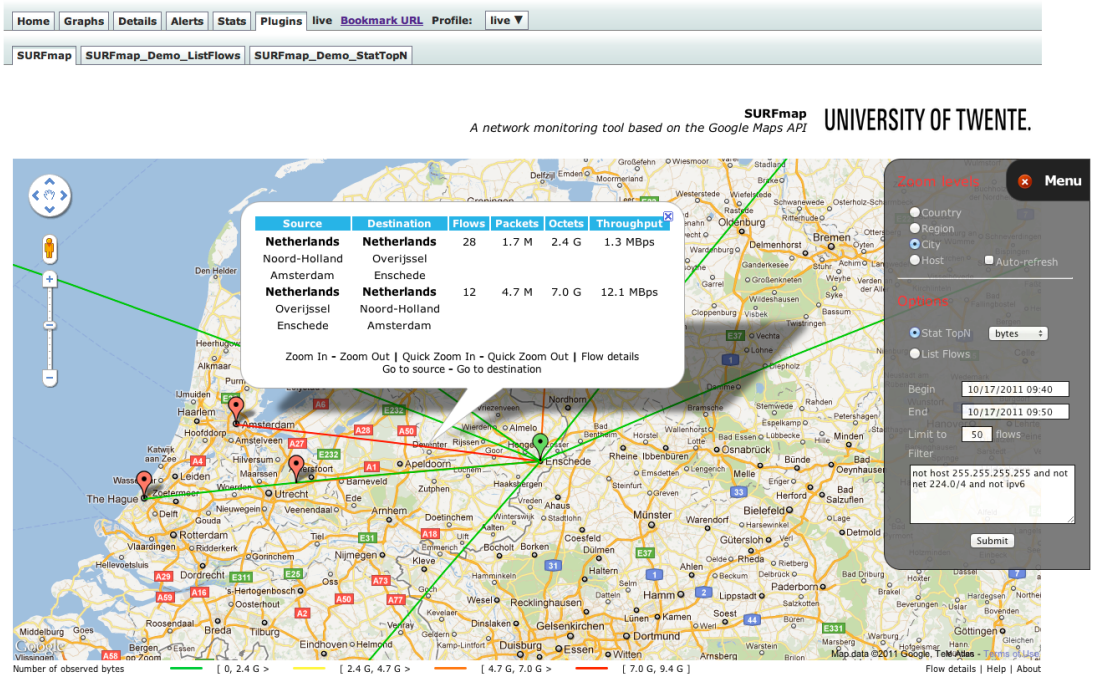

Figure 3: *Stat TopN* at the *City* zoom level

The last screenshot to be discussed in this document is shown in Figure 3. It shows SURFmap at the *Region* zoom level, but it is in NfSen's *Stat TopN* option this time. The criterion to base the statistics on, is set to *bytes*. As a consequence, SURFmap shows the top 50 (so N = 50) of largest flows in the selected time window, based on *bytes*. According to the legend below the map, the selected red line should represent flows which have a total octet sum between 7.0GB and 9.4GB.

The information window shows that a total of 28 flows have been exchanged between the Dutch regions Noord-Holland and Overijssel. In the 12 flows from Overijssel in the direction of Noord-Holland, roughly 4.7GB of network data has been exchanged. In the other direction, however, much less data has been transmitted.

5 Troubleshooting

If you encounter any problems with SURFmap, please perform the following steps:

1. Clear the cache of your Web browser and restart your browser.
2. In case you have installed some new PHP modules (*e.g.* as part of the automated installation procedure by the provided installations scripts), please restart your Web server daemon.
3. Make sure you run SURFmap from within NfSen, instead of as a standalone application. This means that you have to load by navigating to the *Plugins* in NfSen.

In case you encounter any errors, inconsistencies, etc. please send an email with details. You can do this for feature requests as well. Please do always provide as much information and details as possible when making a support request. You can do that, for example, by enabling debug logging[1] and send the resulting output. Your support is appreciated!

E-mail: `r.j.hofstede@utwente.nl`
Mailing list: `https://lists.sourceforge.net/lists/listinfo/surfmap-discuss`

[1] Debug logging is written to syslog and can be enabled in *config.php*.

License

The SURFmap project is distributed under the BSD license:

Copyright (c) 2013, Rick Hofstede (University of Twente, The Netherlands) All rights reserved. Redistribution and use in source and binary forms, with or without modification, are permitted provided that the following conditions are met:

- Redistributions of source code must retain the above copyright notice, this list of conditions and the following disclaimer.
- Redistributions in binary form must reproduce the above copyright notice, this list of conditions and the following disclaimer in the documentation and/or other materials provided with the distribution.
- Neither the name of Rick Hofstede, nor the name of the University of Twente, nor the names of its contributors may be used to endorse or promote products derived from this software without specific prior written permission.

THIS SOFTWARE IS PROVIDED BY THE COPYRIGHT HOLDERS AND CONTRIBUTORS "AS IS" AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE COPYRIGHT OWNER OR CONTRIBUTORS BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

Acknowledgements

This work has been supported by the EC IST-EMANICS Network of Excellence (#26854) and FLAMINGO, a Network of Excellence project (ICT-318488) supported by the European Commission under its Seventh Framework Programme. Special thanks to Pavel Celeda from INVEA-TECH for his valuable contributions.

References

[1] B. Claise, "Cisco Systems NetFlow Services Export Version 9." RFC 3954 (Informational), October 2004.

[2] G. Sadasivan, N. Brownlee, B. Claise, and J. Quittek, "Architecture for IP Flow Information Export." RFC 5470 (Informational), March 2009.

[3] Peter Haag, "NfSen." http://nfsen.sourceforge.net/, 2011. Accessed 7 May 2013.

[4] IP2Location, "IP2Location." http://www.ip2location.com/, 2011. Accessed 7 May 2013.

[5] IP2Location, "MaxMind - GeoLite City." http://www.maxmind.com/app/geolitecity, 2011. Accessed 7 May 2013.

[6] Peter Haag, "NFDUMP." http://nfdump.sourceforge.net/, 2011. Accessed 7 May 2013.